

MOMENT INEQUALITIES OF THE SECOND AND THIRD ORDER

SLAVKO SIMIĆ

ABSTRACT. In this paper we give refinements of some convex and log-convex moment inequalities of the second and third order using a special kind of an algebraic positive semi-definite form. An open problem concerning eight parameter refinement of the third order is also stated with some applications in Information Theory concerning relative divergence of type s .

1. INTRODUCTION

Inequalities for moments of s -th order EX^s , of a probability law with support on \mathbb{R}^+ , are of fundamental interest in Probability Theory. Most known of them all is Jensen's Moment Inequality, given by

$$\begin{aligned} EX^s &\geq (EX)^s, & s \in (-\infty, 0) \cup (1, +\infty); \\ EX^s &\leq (EX)^s, & s \in (0, 1). \end{aligned}$$

Or, more generally,

Theorem A *If F is a convex function on an interval $I \subset \mathbb{R}$, then*

$$E(F(X)) \geq F(EX)$$

for any probability law with support on I .

The topic of research in this article is a difference of moments, introduced in the following way.

For an arbitrary probability law with support on $(0, \infty)$, define

$$\lambda_s = \lambda_s(X) := \begin{cases} (EX^s - (EX)^s)/s(s-1), & s \neq 0, 1; \\ \log(EX) - E(\log X), & s = 0; \\ E(X \log X) - (EX) \log(EX), & s = 1. \end{cases}$$

Throughout the paper we suppose that moments exist for all $s \in I \subset \mathbb{R}$.

The above Jensen's Moment Inequality yields that $\lambda_s \geq 0$ for $s \in I$.

It is also known ([3]) that λ_s is log-convex (hence convex) in s , that is

$$\xi(s, t) := \lambda_s \lambda_t - \lambda_{\frac{s+t}{2}}^2 \geq 0; s, t \in I. \quad (1)$$

Moreover, we proved the following mixture of Jensen-Lyapunov moment inequalities ([4]),

Theorem B *For any $-\infty < r < s < t < +\infty$, we have*

$$(\lambda_s)^{t-r} \leq (\lambda_r)^{t-s} (\lambda_t)^{s-r}. \quad (2)$$

Note that for $s \in \mathbf{N}; r, t \in 2\mathbf{N}$, the relation (2) is valid for arbitrary probability distributions with support on $(-\infty, +\infty)$, ([5]).

2. RESULTS

Our aim in this article is to give some refinements of the above moment inequalities.

Namely, the fact that λ_s is convex i.e. $\tau(s, t) := \lambda_s - 2\lambda_{\frac{s+t}{2}} + \lambda_t \geq 0$, is improved to the following four parameter inequality of second order.

Theorem 1 *For arbitrary $s, t, u, v \in I$, we have*

$$\tau(s, t)\tau(u, v) \geq [\tau(\frac{s+u}{2}, \frac{t+v}{2}) - \tau(\frac{s+v}{2}, \frac{t+u}{2})]^2.$$

As a corollary, we get the next interesting assertion.

Corollary 1 *For fixed $a \in \mathbb{R}$, the function $\mu_a(t) := \lambda_t - 2\lambda_{t+a/2} + \lambda_{t+a}$ is log-convex in t .*

Also,

Theorem 2 *For arbitrary $p, q \geq 0$, the function $w(s, t) := p^2\lambda_s + 2pq\lambda_{\frac{s+t}{2}} + q^2\lambda_t$ is log-convex in $s, t \in I$, that is*

$$w(s, t)w(u, v) \geq [w(\frac{s+u}{2}, \frac{t+v}{2})]^2.$$

The inequality (1) can be improved to the next assertion of second order.

Theorem 3 *Denote*

$$\phi(s, t; u, v) := \xi(s, t)\xi(u, v) - [\xi(\frac{s+u}{2}, \frac{t+v}{2}) - \xi(\frac{s+v}{2}, \frac{t+u}{2})]^2.$$

Then the inequality

$$\phi(s, t; u, v) \geq 0,$$

holds for each $s, t, u, v \in I$.

The above inequality is very sharp. For instance, taking the discrete probability law with masses x and y and assigned probabilities p and q , a calculation shows that

$$\phi(2, 4; 2, 6) = C(p, q)(x - y)^{14},$$

where

$$C(p, q) = \frac{(pq)^4}{5529600} [35 + 11(p - q)^2 + 17(p - q)^4 + (p - q)^6].$$

As a consequence we obtain the next log-convexity assertion.

Corollary 2 *For some $a \in \mathbb{R}$, the function $\sigma(x) = \sigma_a(x) := \xi(x, x + a)$ is log-convex on I .*

Based on a plenty of calculated examples, we conclude that refinements of the third order are also possible.

Its general form is given by the following 8-parameters hypothesis.

Conjecture 1 *The inequality*

$$\begin{aligned} & \phi(r_1, s_1; u_1, v_1) \phi(r_2, s_2; u_2, v_2) \geq \\ & [(\xi(\frac{r_1 + r_2}{2}, \frac{s_1 + s_2}{2}) - \xi(\frac{r_1 + s_2}{2}, \frac{s_1 + r_2}{2}))(\xi(\frac{u_1 + u_2}{2}, \frac{v_1 + v_2}{2}) - \xi(\frac{u_1 + v_2}{2}, \frac{v_1 + u_2}{2})) \\ & - (\xi(\frac{r_1 + u_2}{2}, \frac{s_1 + v_2}{2}) - \xi(\frac{r_1 + v_2}{2}, \frac{s_1 + u_2}{2}))(\xi(\frac{u_1 + r_2}{2}, \frac{v_1 + s_2}{2}) - \xi(\frac{u_1 + s_2}{2}, \frac{v_1 + r_2}{2}))]^2 \\ & \text{holds for arbitrary } r_i, s_i, u_i, v_i \in I, i \in \{1, 2\} ? \end{aligned}$$

We are able to prove Conjecture 1 in some particular cases.

The first one is for $r_1 = u_1 = r_2, s_1 = u_2, v_1 = s_2 = v_2$. Therefore, we obtain a 3-parameter refinement of the third order, given by

Theorem 4 *For any $r, s, v \in I$, we have*

$$\begin{aligned} & \phi(r, s; r, v) \phi(r, v; s, v) \geq \\ & \geq [\xi(r, v)(\xi(s, \frac{r + v}{2}) - \xi(\frac{r + s}{2}, \frac{s + v}{2})) \\ & + (\xi(r, \frac{s + v}{2}) - \xi(\frac{r + s}{2}, \frac{r + v}{2}))(\xi(v, \frac{r + s}{2}) - \xi(\frac{r + v}{2}, \frac{s + v}{2}))]^2. \end{aligned}$$

The second case is conditioned by $r_1 = r_2, s_1 = c_1 = s_2 = u_2$. The 4-parameter improvement of the third order follows as

Theorem 5 *For any $r, s, u, v \in I$, we have*

$$\begin{aligned} & \phi(r, s; s, u) \phi(r, s; s, v) \geq \\ & \geq [\xi(r, s)(\xi(s, \frac{u + v}{2}) - \xi(\frac{s + u}{2}, \frac{s + v}{2})) \\ & - (\xi(s, \frac{r + u}{2}) - \xi(\frac{r + s}{2}, \frac{s + u}{2}))(\xi(s, \frac{r + v}{2}) - \xi(\frac{r + s}{2}, \frac{s + v}{2}))]^2. \end{aligned}$$

Another conjecture is related to the function $\sigma_a(x) := \xi(x, x+a)$. By the result of Corollary 2, we have that

$$\theta_a(x, y) := \sigma_a(x)\sigma_a(y) - \sigma_a^2\left(\frac{x+y}{2}\right) \geq 0.$$

It seems that an inequality analogous to the one from Theorem 3 holds in this case.

Conjecture 2 *Is it true that*

$$\theta_a(x, y)\theta_a(z, v) \geq [\theta_a\left(\frac{x+z}{2}, \frac{y+v}{2}\right) - \theta_a\left(\frac{x+v}{2}, \frac{y+z}{2}\right)]^2?$$

3. PROOFS

In order to prove our results, the following Main Lemma is of crucial importance.

Lemma 3.1. *For arbitrary real numbers a, b, c, d and $r, s, u, v \in I$, the form*

$$\begin{aligned} \psi(a, b, c, d) := & [a^2\lambda_r + 2ab\lambda_{\frac{r+s}{2}} + b^2\lambda_s][c^2\lambda_u + 2cd\lambda_{\frac{u+v}{2}} + d^2\lambda_v] \\ & - [ac\lambda_{\frac{r+u}{2}} + ad\lambda_{\frac{r+v}{2}} + bc\lambda_{\frac{s+u}{2}} + bd\lambda_{\frac{s+v}{2}}]^2 \end{aligned}$$

is positive semi-definite i.e., $\psi(a, b, c, d) \geq 0$.

Proof. For a variable $x \in \mathbb{R}^+$ and a parameter $s \in \mathbb{R}$ consider the function $f_s(x)$, defined as

$$f_s(x) := \begin{cases} (x^s - sx + s - 1)/s(s-1) & , s(s-1) \neq 0; \\ x - \log x - 1 & , s = 0; \\ x \log x - x + 1 & , s = 1. \end{cases}$$

Since,

$$f'_s(x) = \begin{cases} \frac{x^{s-1}-1}{s-1} & , s(s-1) \neq 0; \\ 1 - \frac{1}{x} & , s = 0; \\ \log x & , s = 1, \end{cases}$$

and

$$f''_s(x) = x^{s-2},$$

it follows that $f_s(x)$ is a twice continuously differentiable convex function.

Now, for arbitrary parameters $X, Y \in \mathbb{R}$, denote

$$\begin{aligned} F(x) := & (a^2f_r(x) + 2abf_{\frac{r+s}{2}}(x) + b^2f_s(x))X^2 + 2(acf_{\frac{r+u}{2}}(x) + adf_{\frac{r+v}{2}}(x) + bcf_{\frac{s+u}{2}}(x) + bdf_{\frac{s+v}{2}}(x))XY \\ & + (c^2f_u(x) + 2cdf_{\frac{u+v}{2}}(x) + d^2f_v(x))Y^2. \end{aligned}$$

Since,

$$\begin{aligned} F''(x) = & \frac{1}{x^2}[(a^2x^r + 2abx^{\frac{r+s}{2}} + b^2x^s)X^2 + 2(acx^{\frac{r+u}{2}} + adx^{\frac{r+v}{2}} + bcx^{\frac{s+u}{2}} + bdx^{\frac{s+v}{2}})XY \\ & + (c^2x^u + 2cdx^{\frac{u+v}{2}} + d^2x^v)Y^2] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{x^2}[(ax^{\frac{r}{2}} + bx^{\frac{s}{2}})^2 X^2 + 2(ax^{\frac{r}{2}} + bx^{\frac{s}{2}})(cx^{\frac{u}{2}} + dx^{\frac{v}{2}})XY + (cx^{\frac{u}{2}} + dx^{\frac{v}{2}})^2 Y^2] \\
&= \frac{1}{x^2}[(ax^{\frac{r}{2}} + bx^{\frac{s}{2}})X + (cx^{\frac{u}{2}} + dx^{\frac{v}{2}})Y]^2,
\end{aligned}$$

we conclude that F is a convex function for $x > 0$.

Applying Theorem A in this case, we obtain

$$\begin{aligned}
&(a^2\lambda_r + 2ab\lambda_{\frac{r+s}{2}} + b^2\lambda_s)X^2 + 2(ac\lambda_{\frac{r+u}{2}} + ad\lambda_{\frac{r+v}{2}} + bc\lambda_{\frac{s+u}{2}} + bd\lambda_{\frac{s+v}{2}})XY \\
&\quad + (c^2\lambda_u + 2cd\lambda_{\frac{u+v}{2}} + d^2\lambda_v)Y^2 \geq 0.
\end{aligned}$$

Since the function λ is log-convex, note that the terms

$$\alpha := a^2\lambda_r + 2ab\lambda_{\frac{r+s}{2}} + b^2\lambda_s$$

and

$$\gamma := c^2\lambda_u + 2cd\lambda_{\frac{u+v}{2}} + d^2\lambda_v$$

are non-negative for arbitrary $a, b, c, d \in \mathbb{R}$.

On the other hand, the form

$$\alpha X^2 + 2\beta XY + \gamma Y^2, \quad \alpha, \gamma \geq 0,$$

is positive semi-definite if and only if $\alpha\gamma - \beta^2 \geq 0$.

The proof of Lemma 3.1 readily follows. □

Proof. of Theorem 1.

Taking $a = c = -b = -d = 1$ in Lemma 3.1, we obtain

$$\begin{aligned}
&[\lambda_s - 2\lambda_{\frac{s+t}{2}} + \lambda_t][\lambda_u - 2\lambda_{\frac{u+v}{2}} + \lambda_v] \\
&\geq [\lambda_{\frac{s+u}{2}} - \lambda_{\frac{s+v}{2}} + \lambda_{\frac{t+v}{2}} - \lambda_{\frac{t+u}{2}}]^2,
\end{aligned}$$

which is equivalent to the result of Theorem 1.

For $s = t + a, v = u + a$, we get

$$\mu_a(t)\mu_a(u) \geq \mu_a^2\left(\frac{t+u}{2}\right),$$

i.e. the assertion from Corollary 1. □

Proof. of Theorem 2.

Since $\lambda(s)$ is a convex function, taking $a = c = p; b = d = q; p, q > 0$ in the above lemma, we get

$$w(s, t)w(u, v) = (p^2\lambda_s + 2pq\lambda_{\frac{s+t}{2}} + q^2\lambda_t)(p^2\lambda_u + 2pq\lambda_{\frac{u+v}{2}} + q^2\lambda_v)$$

$$\begin{aligned}
&\geq [p^2\lambda_{\frac{s+u}{2}} + pq(\lambda_{\frac{s+v}{2}} + \lambda_{\frac{t+u}{2}}) + q^2\lambda_{\frac{t+v}{2}}]^2 \\
&\geq [p^2\lambda_{\frac{s+u}{2}} + 2pq\lambda_{\frac{s+u+t+v}{4}} + q^2\lambda_{\frac{t+v}{2}}]^2 = w^2(\frac{s+u}{2}, \frac{t+v}{2}),
\end{aligned}$$

as desired. □

Proof. of Theorem 3.

For complete proof of the assertion of this theorem, we should preliminary prove that it is valid in the special case $v = s$, that is

Lemma 3.2. *For arbitrary $s, t, u \in I$, we have that*

$$\phi(s, t; s, u) := \xi(s, t)\xi(s, u) - [\xi(s, \frac{t+u}{2}) - \xi(\frac{s+u}{2}, \frac{s+t}{2})]^2 \geq 0.$$

Proof. Since

$$\begin{aligned}
\psi(0, 1, c, d) &= \lambda_s[c^2\lambda_u + 2cd\lambda_{\frac{u+v}{2}} + d^2\lambda_v] - [c\lambda_{\frac{s+u}{2}} + d\lambda_{\frac{s+v}{2}}]^2 \\
&= [\lambda(s)\lambda(u) - \lambda^2(\frac{s+u}{2})]c^2 + 2[\lambda(s)\lambda(\frac{u+v}{2}) - \lambda(\frac{s+u}{2})\lambda(\frac{s+v}{2})]cd \\
&\quad + [\lambda(s)\lambda(v) - \lambda^2(\frac{s+v}{2})]d^2,
\end{aligned}$$

and this quadratic form is positive semi-definite, then necessarily,

$$\begin{aligned}
0 &\leq [\lambda_s\lambda_u - \lambda_{\frac{s+u}{2}}^2][\lambda_s\lambda_v - \lambda_{\frac{s+v}{2}}^2] \\
&\quad - [\lambda_s\lambda_{\frac{u+v}{2}} - \lambda_{\frac{s+u}{2}}\lambda_{\frac{s+v}{2}}]^2 = \phi(s, u; s, v).
\end{aligned}$$

□

Choose now numbers a, b, c, d in Lemma 3.1 such that

$$\begin{aligned}
a^2 &= [\lambda_u\lambda_{\frac{s+v}{2}}^2 - 2\lambda_{\frac{s+u}{2}}\lambda_{\frac{s+v}{2}}\lambda_{\frac{u+v}{2}} + \lambda_v\lambda_{\frac{s+u}{2}}^2]/\xi(u, v) \\
&= \frac{1}{\lambda_u}[\lambda_{\frac{s+u}{2}}^2 + (\xi(u, \frac{s+v}{2}) - \xi(\frac{s+u}{2}, \frac{u+v}{2}))^2/\xi(u, v)]; \\
(b\lambda_s + a\lambda_{\frac{r+s}{2}})^2 &= \frac{\xi(r, s)}{\lambda(u)\xi(u, v)}\phi(u, s; u, v); \\
c &= -\lambda_{\frac{s+v}{2}}/a; \quad d = \lambda_{\frac{s+u}{2}}/a.
\end{aligned}$$

A calculation shows that this choice gives

$$\begin{aligned}
a^2\lambda_r + 2ab\lambda_{\frac{r+s}{2}} + b^2\lambda_s &= \frac{1}{\lambda_s}[(b\lambda_s + a\lambda_{\frac{r+s}{2}})^2 + a^2\xi(r, s)] \\
&= \frac{1}{\lambda_s}[\frac{\xi(r, s)}{\lambda_u\xi(u, v)}\phi(u, s; u, v) + \frac{\xi(r, s)}{\lambda_u}[\lambda_{\frac{s+u}{2}}^2 + \frac{(\xi(u, \frac{s+v}{2}) - \xi(\frac{s+u}{2}, \frac{u+v}{2}))^2}{\xi(u, v)}]] \\
&= \frac{\xi(r, s)}{\lambda_u\lambda_s\xi(u, v)}[\phi(u, s; u, v) + (\xi(u, \frac{s+v}{2}) - \xi(\frac{s+u}{2}, \frac{u+v}{2}))^2 + \xi(u, v)\lambda_{\frac{s+u}{2}}^2]
\end{aligned}$$

$$= \frac{\xi(r, s)}{\lambda_u \lambda_s} [\xi(s, u) + \lambda_{\frac{s+u}{2}}^2] = \xi(r, s).$$

By the definition of c and d , we also get

$$\begin{aligned} & c^2 \lambda_u + 2cd \lambda_{\frac{u+v}{2}} + d^2 \lambda_v \\ &= \frac{1}{a^2} [\lambda_u \lambda_{\frac{s+v}{2}}^2 - 2\lambda_{\frac{s+u}{2}} \lambda_{\frac{s+v}{2}} \lambda_{\frac{u+v}{2}} + \lambda_v \lambda_{\frac{s+u}{2}}^2] = \xi(u, v), \end{aligned}$$

and

$$\begin{aligned} & ac \lambda_{\frac{r+u}{2}} + ad \lambda_{\frac{r+v}{2}} + bc \lambda_{\frac{s+u}{2}} + bd \lambda_{\frac{s+v}{2}} \\ &= -\lambda_{\frac{r+u}{2}} \lambda_{\frac{s+v}{2}} + \lambda_{\frac{s+u}{2}} \lambda_{\frac{r+v}{2}} - (b/a) \lambda_{\frac{s+v}{2}} \lambda_{\frac{s+u}{2}} + (b/a) \lambda_{\frac{s+v}{2}} \lambda_{\frac{s+u}{2}} \\ &= \lambda_{\frac{s+u}{2}} \lambda_{\frac{r+v}{2}} - \lambda_{\frac{r+u}{2}} \lambda_{\frac{s+v}{2}} = \xi\left(\frac{s+u}{2}, \frac{r+v}{2}\right) - \xi\left(\frac{r+u}{2}, \frac{s+v}{2}\right). \end{aligned}$$

Therefore, applying Lemma 3.1 for the given choice of parameters, we obtain the result of Theorem 3. □

Proof. of Corollary 2.

Indeed, by Theorem 3 we get

$$\begin{aligned} 0 \leq \phi(x, x+a; y, y+a) &= \xi(x, x+a) \xi(y, y+a) - \left[\xi\left(\frac{x+y}{2}, \frac{x+y}{2} + a\right) - \xi\left(\frac{x+y+a}{2}, \frac{x+y+a}{2}\right) \right]^2 \\ &= \sigma_a(x) \sigma_a(y) - \sigma_a^2\left(\frac{x+y}{2}\right). \end{aligned}$$

□

Proof. of Theorem 4.

In the sequel we need the following elementary assertion.

Lemma 3.3. Denote $D := \alpha\beta - \gamma^2$, $E := \eta - \frac{1}{\alpha}[\delta^2 + \frac{(\alpha\varepsilon - \gamma\delta)^2}{D}]$.

Then the form

$$F(a, c) := \alpha a^2 + \beta c^2 + 2\gamma ac + 2\delta a + 2\varepsilon c + \eta$$

is positive semi-definite if and only if $\alpha \geq 0, D \geq 0, E \geq 0$.

Proof. From the identity

$$F(a, c) = \frac{1}{\alpha} [(\alpha a + \gamma c + \delta)^2 + \frac{1}{D} (Dc + \alpha\varepsilon - \gamma\delta)^2] + E,$$

the proof easily follows. □

Now, developing the form $\psi(a, b, c, d)$ with $b = d = 1, u = r$, we get

$$\begin{aligned} 0 \leq \psi(a, 1, c, 1) &= \xi(r, v)a^2 - 2[\lambda_r \lambda_{\frac{s+v}{2}} - \lambda_{\frac{r+s}{2}} \lambda_{\frac{r+v}{2}}]ac + \xi(r, s)c^2 \\ &\quad + 2[\lambda_v \lambda_{\frac{r+s}{2}} - \lambda_{\frac{v+s}{2}} \lambda_{\frac{r+v}{2}}]a + 2[\lambda_s \lambda_{\frac{r+v}{2}} - \lambda_{\frac{r+s}{2}} \lambda_{\frac{s+v}{2}}]c + \xi(s, v), \end{aligned}$$

and, applying Lemma 3.3 with

$$\begin{aligned} \alpha &= \xi(r, v); \quad \beta = \xi(r, s); \quad \gamma = -[\xi(r, \frac{s+v}{2}) - \xi(\frac{r+s}{2}, \frac{r+v}{2})]; \\ \delta &= [\xi(v, \frac{r+s}{2}) - \xi(\frac{v+s}{2}, \frac{r+v}{2})]; \quad \varepsilon = [\xi(s, \frac{r+v}{2}) - \xi(\frac{r+s}{2}, \frac{s+v}{2})]; \quad \eta = \xi(s, v), \end{aligned}$$

we obtain the proof since in this case

$$D = \phi(r, v; r, s); E = \frac{F}{\xi(r, v)\phi(r, v; r, s)},$$

where F is exactly

$$\begin{aligned} F &= \phi(r, s; r, v)\phi(r, v; s, v) \\ &\quad - [\xi(r, v)(\xi(s, \frac{r+v}{2}) - \xi(\frac{r+s}{2}, \frac{s+v}{2})) \\ &\quad + (\xi(r, \frac{s+v}{2}) - \xi(\frac{r+s}{2}, \frac{r+v}{2}))(\xi(v, \frac{r+s}{2}) - \xi(\frac{r+v}{2}, \frac{s+v}{2}))]^2, \end{aligned}$$

as desired. □

Remark 1 *The side result $D \geq 0$ yields another proof of Lemma 3.2.*

Proof. of Theorem 5.

Developing the form $\phi(a, 1, c, d)$ in a , we get

$$0 \leq \phi(a, 1, c, d) = a^2[\lambda_r C - D^2] + 2a[\lambda_{\frac{r+s}{2}} C - DE] + [\lambda_s C - E^2],$$

where

$$\begin{aligned} C &:= \lambda_u c^2 + 2\lambda_{\frac{u+v}{2}} cd + \lambda_v d^2; \\ D &:= \lambda_{\frac{r+u}{2}} c + \lambda_{\frac{r+v}{2}} d; \\ E &:= \lambda_{\frac{s+u}{2}} c + \lambda_{\frac{s+v}{2}} d. \end{aligned}$$

Hence,

$$[\lambda_r C - D^2][\lambda_s C - E^2] - [\lambda_{\frac{r+s}{2}} C - DE]^2 \geq 0,$$

which is equivalent to

$$\begin{aligned} 0 &\leq \xi(r, s)C + 2\lambda_{\frac{r+s}{2}} DE - \lambda_r E^2 - \lambda_s D^2 \\ &= \alpha c^2 + 2\gamma cd + \beta d^2. \end{aligned}$$

Calculating the coefficients in this case, we obtain

$$\begin{aligned}
\alpha &= \frac{1}{\lambda_s} [\xi(r, s) \xi(u, s) - (\lambda_s \lambda_{\frac{r+u}{2}} - \lambda_{\frac{r+s}{2}} \lambda_{\frac{s+u}{2}})^2] = \frac{\phi(r, s; u, s)}{\lambda_s}; \\
\beta &= \frac{1}{\lambda_s} [\xi(r, s) \xi(v, s) - (\lambda_s \lambda_{\frac{r+v}{2}} - \lambda_{\frac{r+s}{2}} \lambda_{\frac{s+v}{2}})^2] = \frac{\phi(r, s; v, s)}{\lambda_s}; \\
\gamma &= \frac{1}{\lambda_s} [\xi(r, s) (\lambda_s \lambda_{\frac{u+v}{2}} - \lambda_{\frac{s+u}{2}} \lambda_{\frac{s+v}{2}}) - (\lambda_s \lambda_{\frac{r+u}{2}} - \lambda_{\frac{r+s}{2}} \lambda_{\frac{s+u}{2}}) (\lambda_s \lambda_{\frac{r+v}{2}} - \lambda_{\frac{r+s}{2}} \lambda_{\frac{s+v}{2}})] \\
&= \frac{1}{\lambda_s} [\xi(r, s) (\xi(s, \frac{u+v}{2}) - \xi(\frac{s+u}{2}, \frac{s+v}{2})) - (\xi(s, \frac{r+u}{2}) - \xi(\frac{r+s}{2}, \frac{s+u}{2})) (\xi(s, \frac{r+v}{2}) - \xi(\frac{r+s}{2}, \frac{s+v}{2}))] \\
&\text{and, since } \alpha\beta - \gamma^2 \geq 0, \text{ the proof follows.}
\end{aligned}$$

□

4. APPLICATIONS IN INFORMATION THEORY

Let

$$\Omega = \{p = \{p_i\} \mid p_i > 0, \sum p_i = 1\},$$

be the set of finite or infinite discrete probability distributions.

One of the most general probability measures which is of importance in Information Theory is the famous Csiszár's f -divergence $C_f(p||q)$ ([5]), defined by

Definition 1 For a convex function $f : (0, \infty) \rightarrow \mathbb{R}$, the f -divergence measure is given by

$$C_f(p||q) := \sum q_i f(p_i/q_i),$$

where $p, q \in \Omega$. By the well known Jensen's inequality for convex functions it follows that

$$C_f(p||q) \geq f(1),$$

with equality if and only if $p = q$.

Some important information measures are just particular cases of the Csiszár's f -divergence.

For example,

(a) taking $f(x) = x^\alpha$, $\alpha > 1$, we obtain the α -order divergence defined by

$$I_\alpha(p||q) := \sum p_i^\alpha q_i^{1-\alpha};$$

Remark 2 The above quantity is an argument in well-known theoretical divergence measures such as Renyi α -order divergence $I_\alpha^R(p||q)$ or Tsallis divergence $I_\alpha^T(p||q)$, defined as

$$I_\alpha^R(p||q) := \frac{1}{\alpha - 1} \log I_\alpha(p||q); \quad I_\alpha^T(p||q) := \frac{1}{\alpha - 1} (I_\alpha(p||q) - 1).$$

(b) for $f(x) = x \log x$, one obtain the Kullback-Leibler divergence ([3]) defined by

$$K(p||q) := \sum p_i \log(p_i/q_i);$$

(c) for $f(x) = (\sqrt{x} - 1)^2$, one obtain the Hellinger distance

$$H^2 = H^2(p, q) = H^2(q, p) := \sum (\sqrt{p_i} - \sqrt{q_i})^2;$$

(d) if we choose $f(x) = (x - 1)^2$, then we get the χ^2 -distance

$$\chi^2(p, q) := \sum (p_i - q_i)^2 / q_i.$$

Besides, the quantity

$$I_{1/2}(p||q) = \sum \sqrt{p_i q_i} := B(p, q) = B(q, p) = B$$

is known as *Bhattacharya coefficient*. Evidently,

$$0 < B \leq 1; \quad 2(1 - B(p, q)) = H^2(p, q).$$

We quote now inequalities between those measures which are already known in the literature ([1], [7]).

$$\begin{aligned} \chi^2(p, q) &\geq H^2(p, q); \\ K(p||q) &\geq H^2(p, q); \\ K(p||q) &\leq \log(1 + \chi^2(p, q)). \end{aligned} \tag{*}$$

In particular, $K(p||q) \leq \chi^2(p, q)$ ([2]).

The generalized measure $K_s(p||q)$, known as *the relative divergence of type s* ([6], [7]), is defined by

$$K_s(p||q) := \begin{cases} (\sum p_i^s q_i^{1-s} - 1) / (s(s-1)) & , s \in \mathbb{R} / \{0, 1\}; \\ K(q||p) & , s = 0; \\ K(p||q) & , s = 1. \end{cases}$$

It include the Hellinger and χ^2 distances as particular cases.

Indeed,

$$\begin{aligned} K_{1/2}(p||q) &= 4(1 - \sum \sqrt{p_i q_i}) = 2 \sum (p_i + q_i - 2\sqrt{p_i q_i}) = 2H^2(p, q); \\ K_2(p||q) &= \frac{1}{2} \left(\sum \frac{p_i^2}{q_i} - 1 \right) = \frac{1}{2} \sum \frac{(p_i - q_i)^2}{q_i} = \frac{1}{2} \chi^2(p, q). \end{aligned}$$

It will be proved next that $K_s(p||q)$ is log-convex in s for $s \in \mathbb{R}$, wherefrom a whole variety of inequalities connecting the above mentioned measures arise.

For an application of our results, we shall consider discrete probability laws in the sequel. The continuous case can be treated analogously.

For $p, q \in \Omega$, let $\mathcal{A}(q, X)$ be the probability law with the set of positive discrete random variables X and assigned probabilities q . Choosing $X = \frac{p}{q}$, we get

$$EX = \sum q_i \left(\frac{p_i}{q_i} \right) = \sum p_i = 1; \quad EX^s = \sum q_i \left(\frac{p_i}{q_i} \right)^s = \sum p_i^s q_i^{1-s} = I_s(p||q).$$

Therefore, by continuity, for the probability law $\mathcal{A}(q, \frac{p}{q})$ we obtain that $\lambda_s = K_s(p||q)$ for each $s \in \mathbb{R}$.

Hence an important consequence follows.

Theorem 6 *The relative divergence of type s is logarithmically convex in $s \in \mathbb{R}$.*

Noting that $\lambda_{1-s} = K_s(q||p)$, we obtain the following universal estimations.

Theorem 7 *For each $s \in \mathbb{R}$, we have*

$$K_s(p||q)K_s(q||p) \geq 4H^4,$$

and

$$K_s(p||q) + K_s(q||p) \geq 4H^2,$$

where H , as usual, denotes the Hellinger distance.

Proof. Indeed,

$$K_s(p||q)K_s(q||p) = \lambda_s \lambda_{1-s} \geq \lambda_{1/2}^2 = K_{1/2}^2 = 4H^4.$$

Furthermore,

$$K_s(p||q) + K_s(q||p) \geq 2\sqrt{K_s(p||q)K_s(q||p)} \geq 4H^2.$$

□

Denoting the symmetric measures S_a and P_a by

$$S_a = S_a(p||q) = S_a(q||p) := K_a(p||q) + K_a(q||p) - 4H^2;$$

$$P_a = P_a(p||q) = P_a(q||p) := K_a(p||q)K_a(q||p) - 4H^4,$$

and applying Theorems 1 and 3, we get

Theorem 8 *Estimations*

$$|S_a - S_b| \leq (S_{a+b-1/2} S_{a-b+1/2})^{1/2};$$

$$|P_a - P_b| \leq (P_{a+b-1/2} P_{a-b+1/2})^{1/2},$$

hold for each $a, b \in \mathbb{R}$.

Further illustration will be given by improving the classical inequalities (*) for Kullback-Leibler divergence $K(p||q)$.

Theorem 9 *We have*

$$f_1(B, \chi^2) \leq K(p||q) \leq f_2(B, \chi^2),$$

where

$$f_1(B, \chi^2) = -2 \log B + 6 \frac{(1 - B^2)^2}{1 - B^4 + \chi^2(q||p)}$$

and

$$f_2(B, \chi^2) = \log(1 + \chi^2(p, q)) - \frac{32}{9} \frac{(B \sqrt{1 + \chi^2(p, q)} - 1)^2}{\chi^2(p, q)}.$$

Proof. Indeed, for the law $\mathcal{A}(p, \frac{p}{q})$ a calculation shows that

$$\begin{aligned} \lambda_0 &= \log \sum \frac{p_i^2}{q_i} - \sum p_i \log \frac{p_i}{q_i} = \log(1 + \chi^2(p, q)) - K(p||q); \\ \lambda_{-1} &= \frac{1}{2} [\sum p_i (p_i/q_i)^{-1} - (\sum p_i^2/q_i)^{-1}] = \frac{1}{2} (1 - \frac{1}{1 + \chi^2(p, q)}); \\ \lambda_{-1/2} &= \frac{4}{3} [\sum p_i (p_i/q_i)^{-1/2} - (\sum p_i^2/q_i)^{-1/2}] = \frac{4}{3} (B(p, q) - \frac{1}{\sqrt{1 + \chi^2(p, q)}}). \end{aligned}$$

Now, since $\lambda_0 \lambda_{-1} \geq \lambda_{-1/2}^2$, we obtain

$$K(p||q) \leq \log(1 + \chi^2(p, q)) - \frac{32}{9} \frac{(B(p, q) \sqrt{1 + \chi^2(p, q)} - 1)^2}{\chi^2(p, q)},$$

which is a considerable improvement of the target inequality in (*).

To improve lower bound of $K(p||q)$ let us consider the law $\mathcal{A}(p, \sqrt{\frac{q}{p}})$. Since then

$$\lambda_0 = \log(\sum p_i \sqrt{q_i/p_i}) - \sum p_i \log(\sqrt{q_i/p_i}) = \log B(p, q) + \frac{1}{2} K(p||q),$$

and $\lambda_0 \geq 0$, we get a better approximation at once

$$K(p||q) \geq -2 \log B(p, q).$$

Indeed,

$$-2 \log B(p, q) = -2 \log(1 - H^2(p, q)/2) \geq H^2(p, q).$$

We can get further improvement by the inequality $\lambda_0 \lambda_4 \geq \lambda_2^2$.

Because,

$$\lambda_2 = \frac{1}{2} (\sum p_i (\sqrt{q_i/p_i})^2 - (\sum p_i \sqrt{q_i/p_i})^2) = \frac{1}{2} (1 - B^2(p, q)),$$

and

$$\lambda_4 = \frac{1}{12} (\sum p_i (\sqrt{q_i/p_i})^4 - (\sum p_i \sqrt{q_i/p_i})^4) = \frac{1}{12} (\chi^2(q||p) + 1 - B^4(p, q)),$$

we finally obtain

$$K(p||q) \geq -2 \log B(p, q) + 6 \frac{(1 - B^2(p, q))^2}{1 - B^4(p, q) + \chi^2(q||p)}.$$

□

In this way we get better approximation of Kullback-Leibler divergence in terms of Bhattacharya coefficient and χ^2 - distance.

REFERENCES

- [1] Cover, T. M., Thomas, J. A., *Elements of Information Theory*, John Wiley and Sons Inc., New York (1991).
- [2] Dragomir, S. S., Gluscevic, V., *Some inequalities for the Kullback-Leibler and χ^2 - distances in Information Theory and applications*, Tamsui Oxford Journal of Mathematical Sciences 17(2) (2001), 97-111.
- [3] Simic, S., *On logarithmic convexity for differences of power means*, J. Inequal. Appl. Art. ID 37359 (2007), 8 pp.
- [4] Simic, S., *On a new moment inequality*, Statist. Probab. Lett. **78**, No. 16 (2008), 2671-2678.
- [5] Simic, S., *On a moment inequality for laws on $(-\infty, +\infty)$* , Statist. Probab. Lett. **79**, No. 22 (2009), 2334-2337.
- [6] Taneja, I. J., *New developments in generalized information measures*, Advances in Imaging and Electron Physics, **91** (1995), 37-135.
- [7] Vajda, I., *Theory of Statistical Inference and Information*, Kluwer Academic Press, London (1989).

MATHEMATICAL INSTITUTE SANU, KNEZA MIHAILA 36, 11000 BELGRADE, SERBIA

E-mail address: `ssimic@turing.mi.sanu.ac.rs`